# Data Science Project Proposal

*Black Friday Data Set*

*: Predict Top 5 product items most likely to be purchased by a customer*

2017320160 임혜수

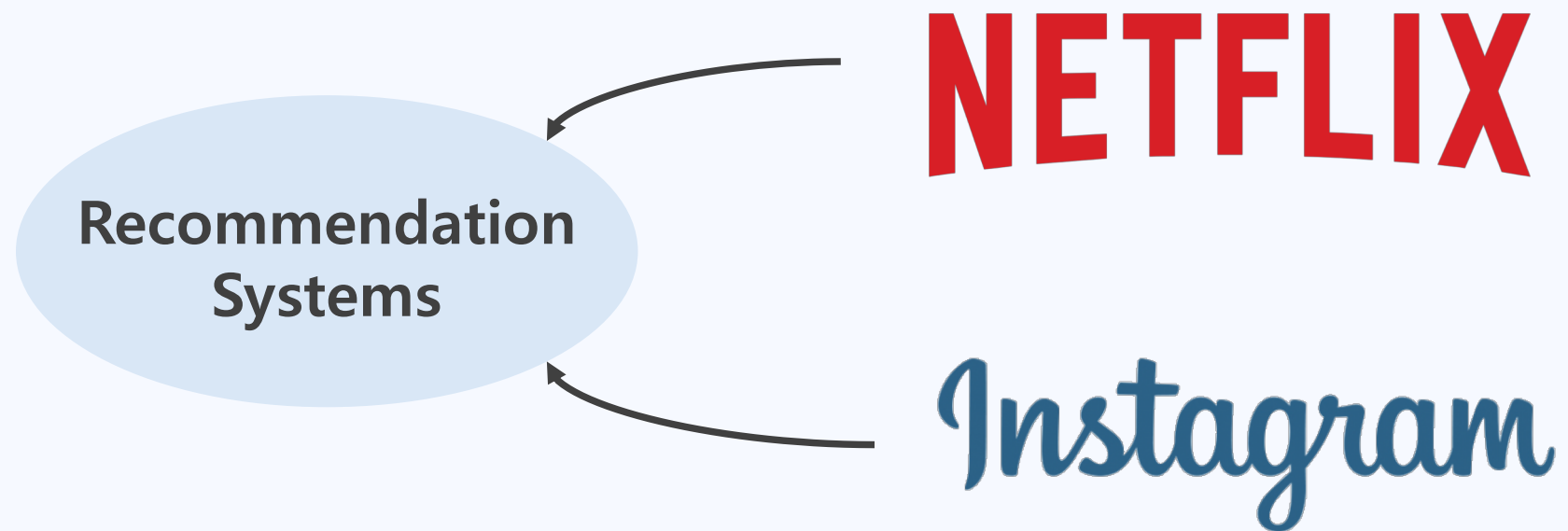2017320214 이지원

2017320233 김채령

# CONTENTS

Recommendation Systems

NETFLIX

Instagram

# DATASET

- Title : Black Friday Dataset

- Contents : a sample of the transaction made in a retail store on a Black Friday

- number of observations : 550,000

- number of attributes : 12

  - ✓ About user : User_ID, Age, Gender, City, Occupation

  - ✓ About product : Product_ID, Category, Price

- Data Format : CSV file

- Missing Value : exists

- Retrieved From : Kaggle

## PROBLEM DEFINITION

- **Predict** top 5 most likely to be purchased products for customers.

- **Analyze correlation** between attributes.

## PROJECT OUTLINE

1. Data Analysis

2. Data Preprocessing

3. Correlation Analysis

4. Build Recommendation System

   - ML model : item- based collaborative filtering

   - Evaluation metric : RMSE, Recall, Precision, F-score

5. Summary

# EXPECTED RESULT

- Goal

  Build **Recommendation system** to predict top 5 products most likely to be purchased by a customer

- Expected Usage

  **For Seller**

  Predict the products likely to be purchased by customers → Predict sales amount

  Correlation between users and products → Use for marketing strategies

  **For Customer**

  Customers can get recommendations of products

## ROLE

- 임혜수 : Data Analysis. Data Interpretation.

- 이지원 : Documentation. Repository Management.

- 김채령 : Model Building. Model Implementation.

# THANK YOU